

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 2002年 9月 3日
Date of Application:

出願番号 特願2002-257545
Application Number:

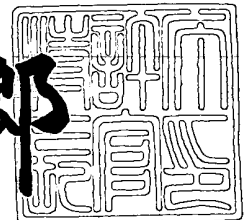
[ST-10/C]: [J.P.2002-257545]

出願人 日本電気株式会社
Applicant(s):

2003年 7月 9日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3055323

【書類名】 特許願

【整理番号】 67000090

【提出日】 平成14年 9月 3日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 13/00

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 酒井 智弘

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100088959

 【弁理士】

 【氏名又は名称】 境 廣巳

【手数料の表示】

 【予納台帳番号】 009715

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

 【包括委任状番号】 9002136

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ループ状インタフェースの障害解析方法及び障害解析機能を有するシステム

【特許請求の範囲】

【請求項 1】 多重化されたループ状インタフェースを有し、前記各ループ状インタフェースへのデバイスの接続および切り離しを行うループ接続切替手段を制御する制御デバイスが前記各ループ状インタフェースに接続され、且つ、前記制御デバイスが相互に通信するためのインタフェースを備えたシステムの障害解析方法であって、前記制御デバイスにおいて全ての前記ループ状インタフェースが異常となったことを検出したときに前記ループ接続切替手段を制御して少なくとも 1 つの前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行うことを特徴とするループ状インタフェースの障害解析方法。

【請求項 2】 前記制御デバイスは、自デバイスが接続されるループ状インタフェースを通じて定期的に送られてくるコマンドの受信が途絶えたことを検出して前記インタフェースを通じて他の前記制御デバイスに通知し、全ての前記制御デバイスにおいて前記コマンドの受信が途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出することを特徴とする請求項 1 記載のループ状インタフェースの障害解析方法。

【請求項 3】 接続された全ての前記デバイスが切り離されてループ異常が解消した前記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施することを特徴とする請求項 1 記載のループ状インタフェースの障害解析方法。

【請求項 4】 多重化されたループ状インタフェースの 1 つのループ状インタフェースを通じて前記デバイスおよび前記制御デバイスに接続された制御装置が、多重化されたループ状インタフェースの全ループ異常時、自制御装置が接続されたループ状インタフェースのループ異常が一定時間内に解消するか否かを判定し、一定時間内にループ異常が解消した場合には前記制御デバイスに全ての前記デバイスの切り離しを行ったか否かを問い合わせ、前記制御デバイスによって全

ての前記デバイスの切り離しが行われていたときに前記ループ診断手段によるループ診断を含む両系リンク障害対策処理を実行することを特徴とする請求項 3 記載のループ状インタフェースの障害解析方法。

【請求項 5】 前記一定時間は、多重化されたループ状インタフェースに全ループ異常が発生した場合に前記制御デバイスが全ての前記デバイスの切り離しを行うのに要する時間より若干長い時間に設定されていることを特徴とする請求項 4 記載のループ状インタフェースの障害解析方法。

【請求項 6】 前記ループ診断で故障と判断された前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させることを特徴とする請求項 3 乃至 5 の何れか 1 項に記載のループ状インタフェースの障害解析方法。

【請求項 7】 利用を再開した前記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施することを特徴とする請求項 6 記載のループ状インタフェースの障害解析方法。

【請求項 8】 前記ループ状インタフェースが FC-AL (F i b r e C h a n n e l A r b i t r a t e d L o o p) であることを特徴とする請求項 1 乃至 7 の何れか 1 項に記載のループ状インタフェースの障害解析方法。

【請求項 9】 前記デバイスがハードディスク装置であり、前記制御デバイスがエンクロージャサービスデバイスであることを特徴とする請求項 1 乃至 7 の何れか 1 項に記載のループ状インタフェースの障害解析方法。

【請求項 1 0】 多重化されたループ状インタフェースを有し、前記各ループ状インタフェースへのデバイスの接続および切り離しを行うループ接続切替手段を制御する機能を有する制御デバイスが前記各ループ状インタフェースに接続されたシステムにおいて、前記制御デバイスは相互に通信するためのインタフェースを備え、且つ、全ての前記ループ状インタフェースが異常となったことを検出したときに前記ループ接続切替手段を制御して少なくとも 1 つの前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行う手段を備えたことを特徴とするループ状インタフェースの障害解析機能を有するシステム

。 【請求項 1 1】 前記制御デバイスは、自デバイスが接続されるループ状インタフェースを通じて定期的に送られてくるコマンドの受信が途絶えたことを検出して前記インタフェースを通じて他の前記制御デバイスに通知する手段を備え、全ての前記制御デバイスにおいて前記コマンドの受信が途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出するものであることを特徴とする請求項 1 0 記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項 1 2】 接続された全ての前記デバイスが切り離されてループ異常が解消した前記ループ状インタフェースに接続された前記制御デバイスを經由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するループ診断手段を備えたことを特徴とする請求項 1 0 記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項 1 3】 多重化されたループ状インタフェースの 1 つのループ状インタフェースを通じて前記デバイスおよび前記制御デバイスに接続された制御装置を備え、該制御装置は、多重化されたループ状インタフェースの全ループ異常時、自制御装置が接続されたループ状インタフェースのループ異常が一定時間内に解消するか否かを判定し、一定時間内にループ異常が解消した場合には前記制御デバイスに全ての前記デバイスの切り離しを行ったか否かを問い合わせ、前記制御デバイスによって全ての前記デバイスの切り離しが行われていたときに前記ループ診断手段によるループ診断を含む両系リンク障害対策処理を実行するものであることを特徴とする請求項 1 2 記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項 1 4】 前記一定時間は、多重化されたループ状インタフェースに全ループ異常が発生した場合に前記制御デバイスが全ての前記デバイスの切り離しを行うのに要する時間より若干長い時間に設定されていることを特徴とする請求項 1 3 記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項 1 5】 前記ループ診断手段は、前記ループ診断で故障と判断され

た前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させるものであることを特徴とする請求項12乃至14の何れか1項に記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項16】 前記ループ診断手段は、利用を再開した前記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するものであることを特徴とする請求項15記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項17】 前記ループ状インタフェースがFC-AL (Fibre Channel Arbitrated Loop) であることを特徴とする請求項10乃至16の何れか1項に記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項18】 前記デバイスがハードディスク装置であり、前記制御デバイスがエンクロージャサービスデバイスであることを特徴とする請求項10乃至16の何れか1項に記載のループ状インタフェースの障害解析機能を有するシステム。

【請求項19】 多重化されたループ状インタフェースの1つのループ状インタフェースに接続され、接続された前記ループ状インタフェースへのデバイスの接続および切り離しを行うループ接続切替手段を制御する機能を有するエンクロージャサービスデバイスにおいて、他のループ状インタフェースに接続された他のエンクロージャサービスデバイスと相互に通信するためのインタフェースを備え、且つ、全ての前記ループ状インタフェースが異常となったことを検出したときに前記ループ接続切替手段を制御して前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行う手段を備えたことを特徴とするエンクロージャサービスデバイス。

【請求項20】 自デバイスが接続されるループ状インタフェースを通じて定期的に送られてくるコマンドが途絶えたことを検出して前記インタフェースを通じて他の前記エンクロージャサービスデバイスに通知する手段を備え、全ての

前記エンクロージャサービスデバイスにおいて前記コマンドが途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出するものであることを特徴とする請求項 19 記載のエンクロージャサービスデバイス。

【請求項 21】 前記ループ状インタフェースが FC-AL (Fibre Channel Arbitrated Loop) であることを特徴とする請求項 19 または 20 に記載のエンクロージャサービスデバイス。

【請求項 22】 前記デバイスがハードディスク装置であることを特徴とする請求項 19 または 20 に記載のエンクロージャサービスデバイス。

【請求項 23】 多重化されたループ状インタフェースの 1 つのループ状インタフェースを通じて、1 以上のデバイスと、前記デバイスのループ状インタフェースへの接続および切り離しを行うループ接続切替手段を制御する機能を有する制御デバイスとに接続された制御装置において、多重化されたループ状インタフェースの全てが異常となったことを検出した前記制御デバイスによって少なくとも 1 つの前記ループ状インタフェースに接続された全ての前記デバイスが切り離されたという事象を確認する手段と、接続された全ての前記デバイスが切り離されてループ異常が解消した前記ループ状インタフェースに接続された前記制御デバイスを経由して、前記制御デバイス間を相互に通信可能に接続するインタフェースを通じて、他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するループ診断手段を備えたことを特徴とする制御装置。

【請求項 24】 前記ループ診断手段は、前記ループ診断で故障と判断された前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させるものであることを特徴とする請求項 23 記載の制御装置。

【請求項 25】 前記ループ診断手段は、利用を再開した前記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するものであることを特徴とする請求項 24 記載の制御装置。

【請求項 26】 前記ループ状インタフェースが FC-AL (Fibre Channel Arbitrated Loop) であることを特徴とする請求項 23 乃至 25 の何れか 1 項に記載の制御装置。

【請求項 27】 前記デバイスがハードディスク装置であり、前記制御デバイスがエンクロージャサービスデバイスであることを特徴とする請求項 23 乃至 25 の何れか 1 項に記載の制御装置。

【請求項 28】 多重化されたループ状インタフェースの 1 つのループ状インタフェースに接続され、接続された前記ループ状インタフェースへのデバイスの接続および切り離しを行うループ接続切替手段を制御する機能を有し、且つ、他のループ状インタフェースに接続された他のエンクロージャサービスデバイスと相互に通信するためのインタフェースを備えたエンクロージャサービスデバイスを構成するコンピュータを、全ての前記ループ状インタフェースが異常となったことを検出する検出手段、該検出手段による検出時に前記ループ接続切替手段を制御して前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行う手段、として機能させるエンクロージャサービスデバイス用プログラム。

【請求項 29】 前記検出手段は、自デバイスが接続されるループ状インタフェースを通じて定期的に送られてくるコマンドが途絶えたことを検出して前記インタフェースを通じて他の前記エンクロージャサービスデバイスに通知する手段を備え、全ての前記エンクロージャサービスデバイスにおいて前記コマンドが途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出するものであることを特徴とする請求項 28 記載のエンクロージャサービスデバイス用プログラム。

【請求項 30】 多重化されたループ状インタフェースの 1 つのループ状インタフェースを通じて、1 以上のデバイスと、前記デバイスのループ状インタフェースへの接続および切り離しを行うループ接続切替手段を制御する機能を有する制御デバイスとに接続された制御装置を構成するコンピュータを、多重化されたループ状インタフェースの全てが異常となったことを検出した前記制御デバイスによって少なくとも 1 つの前記ループ状インタフェースに接続された全ての前

記デバイスが切り離されたという事象を確認する手段と、接続された全ての前記デバイスが切り離されてループ異常が解消した前記ループ状インタフェースに接続された前記制御デバイスを経由して、前記制御デバイス間を相互に通信可能に接続するインタフェースを通じて、他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するループ診断手段、として機能させる制御装置用プログラム。

【請求項 3 1】 前記ループ診断手段は、前記ループ診断で故障と判断された前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させるものであることを特徴とする請求項 3 0 記載の制御装置用プログラム。

【請求項 3 2】 前記ループ診断手段は、利用を再開した前記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するものであることを特徴とする請求項 3 1 記載の制御装置用プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はループ状インタフェースの障害解析方法及びループ状インタフェースの障害解析機能を有するシステムに関する。

【0002】

【従来の技術】

FC-AL (Fibre Channel Arbitrated Loop) のようなループ状にデバイスが接続されるインタフェースは、ケーブルの接続構成がシンプルでデバイスの増設等への対応も容易なことから、ディスクアレイ装置などに近年多く使用されるようになってきた。しかし、このような形態のインタフェースは、接続されているデバイスのインタフェース回路の障害等によってループ内を正常に信号が伝搬できなくなると（これを、ループ異常あるいはリンクダウンなどと呼ぶ）、ループ全体が使用不可能になり、接続されている 1 デバ

イスの障害であっても、そのループに接続されている全てのデバイスが使用できなくなる。このため、通常これらのデバイスは独立した2つのループに接続できるようにインタフェース回路を2ポート分備え、2重化されたループ状インタフェースの片方のループが障害により使用できなくなっても、もう一方のループでアクセス可能なようにして信頼性を高めている。また、ループ異常が発生した場合、速やかに故障デバイスを特定して障害を回復させる機能を備えている。

【0 0 0 3】

例えば、特許文献1では、ループA、Bの2つのFC-A Lを有するディスクアレイ装置において、一方のループAに異常が発生したとき、そのループAのディスク制御部は通常のディスク処理を停止して仮縮退状態へ遷移し、ループBのディスク制御部はループAの処理を代行する。そして、ループAのディスク制御部は自らループAの診断を実行して故障ディスクをループAからバイパスし、ループAの仮縮退状態を解除する。これによって、ループ異常が発生した場合にホストへの応答を継続しながら故障部品を特定する処理を可能にしている。但し、この特許文献1では、ループ異常として間欠障害を想定しており、ループAの診断時にはループAが異常でないことを前提としている。

【0 0 0 4】

また、特許文献2には、ループA、Bの2つのFC-A Lを有し且つ各ループにエンクロージャサービスデバイス (E n c l o s u r e S e r v i c e D e v i c e) が直接接続されたディスクアレイ装置に適用可能な障害解析方法が開示されている。エンクロージャサービスデバイスはANSI規格 (NCITS 305-199X) で規定されている装置の筐体内環境監視・資源管理のためのデバイスであり、自身が接続されるループに接続されているディスクのバイパス、バイパス解除を行うループ接続切替部に対する制御機能も有している。ループに直接接続されたエンクロージャサービスデバイスは、ループに接続されたディスクと同様にループ状インタフェースに接続されたデバイスとして振る舞うため、接続されたループに固定的な異常が発生すると、そのループのディスク制御部からそのループを通じてエンクロージャサービスデバイスに指示を与えることができなくなる。そこで特許文献2では、各ループに接続されているエンクロージャ

ャサービスデバイスに相互に通信するためのインタフェースを設け、一方のループAに異常が発生したとき、そのループAの診断を行うためのループAのディスクのバイパス、バイパス解除の指示を、他方のループBのエンクロージャサービスデバイス経由でループAのエンクロージャサービスデバイスに与えるようにしている。

【0 0 0 5】

【特許文献1】

特開 2 0 0 2 - 7 0 7 7 号公報

【特許文献2】

特開 2 0 0 1 - 2 1 6 2 0 6 号公報

【0 0 0 6】

【発明が解決しようとする課題】

上述したようにF C - A L等のループ状インタフェースを有するディスクアレイ装置などのシステムでは、ループ状インタフェースの2重化、ループ異常時における速やかな障害解析による故障部品のバイパス化を行うことで信頼性を高めるようにしている。しかし、2重化されたループ状インタフェースの双方が同時にループ障害になった場合、各ループにエンクロージャサービスデバイスを直接接続したシステムにあっては、以下のような課題がある。

【0 0 0 7】

第1の課題は、ループ状インタフェースに接続されたデバイスの切り離し、再接続が全く行えなくなることである。その理由は、デバイスの切り離し、再接続を実際に行うループ接続切替部を制御するエンクロージャサービスデバイス自体にどのループからもアクセスすることができなくなるからである。

【0 0 0 8】

第2の課題は、ループ異常の原因となる故障デバイスを特定するループ診断が全く行えないことである。その理由は、ループ診断を実施するためには、ループ状インタフェースに接続されたデバイスの切り離し、再接続が必要であるが、このデバイスの切り離し、再接続が行えないからである。

【0 0 0 9】

第3の課題は、システムが完全に停止したままになることである。その理由はループ診断が実施できず、故障部品をループから除去できないからである。

【0010】

本発明はこのような事情に鑑みて提案されたものであり、その目的は、多重化されたループ状インタフェースを有し且つその各々のループにエンクロージャサービスデバイス等の制御デバイスが接続されたディスクアレイ装置などのシステムにおいて、全てのループで異常が発生しても、ループ状インタフェースに接続されたデバイスの切り離し、再接続が行えるようにし、また、ループ異常の原因となる故障デバイスを特定するループ診断を実施できるようにし、また、システムが完全に停止したままになるのを防止することにある。

【0011】

【課題を解決するための手段】

本発明のループ状インタフェースの障害解析方法は、多重化されたFC-A L等のループ状インタフェースを有し、前記各ループ状インタフェースへのハードディスク装置等のデバイスの接続および切り離しを行うループ接続切替手段を制御するエンクロージャサービスデバイス等の制御デバイスが前記各ループ状インタフェースに接続され、且つ、前記制御デバイスが相互に通信するためのインタフェースを備えたディスクアレイ装置等のシステムの障害解析方法であって、前記制御デバイスにおいて全ての前記ループ状インタフェースが異常となったことを検出したときに前記ループ接続切替手段を制御して少なくとも1つの前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行うことを特徴とする。

【0012】

ここで、前記制御デバイスは、自デバイスが接続されるループ状インタフェースを通じて定期的に送られてくるハートビートコマンド等のコマンドが途絶えたことを検出して前記インタフェースを通じて他の前記制御デバイスに通知し、全ての前記制御デバイスにおいて前記コマンドが途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出するようにして良い。また、接続された全ての前記デバイスが切り離されてループ異常が解消した前

記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するようにして良い。更に、多重化されたループ状インタフェースの1つのループ状インタフェースを通じて前記デバイスおよび前記制御デバイスに接続された制御装置が、多重化されたループ状インタフェースの全ループ異常時、自制御装置が接続されたループ状インタフェースのループ異常が一定時間内に解消するか否かを判定し、一定時間内にループ異常が解消した場合には前記制御デバイスに全ての前記デバイスの切り離しを行ったか否かを問い合わせ、前記制御デバイスによって全ての前記デバイスの切り離しが行われていたときに前記ループ診断手段によるループ診断を含む両系リンク障害対策処理を実行するようにして良い。また、前記ループ診断で故障と判断された前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させるようにして良い。また、利用を再開した前記ループ状インタフェースに接続された前記制御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するようにして良い。

【0013】

本発明のループ状インタフェースの障害解析機能を有するシステムは、多重化されたFC-A L等のループ状インタフェースを有し、前記各ループ状インタフェースへのハードディスク装置等のデバイスの接続および切り離しを行うループ接続切替手段を制御する機能を有するエンクロージャサービスデバイス等の制御デバイスが前記各ループ状インタフェースに接続されたディスクアレイ装置等のシステムにおいて、前記制御デバイスは相互に通信するためのインタフェースを備え、且つ、全ての前記ループ状インタフェースが異常となったことを検出したときに前記ループ接続切替手段を制御して少なくとも1つの前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行う手段を備えたことを特徴とする。

【0014】

ここで、前記制御デバイスは、自デバイスが接続されるループ状インタフェース

を通じて定期的に送られてくるハートビートコマンド等のコマンドが途絶えたことを検出して前記インタフェースを通じて他の前記制御デバイスに通知する手段を備え、全ての前記制御デバイスにおいて前記コマンドが途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出するものであって良い。また、接続された全ての前記デバイスが切り離されてループ異常が解消した前記ループ状インタフェースに接続された前記制御デバイスを經由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するループ診断手段を備えていて良い。更に、多重化されたループ状インタフェースの1つのループ状インタフェースを通じて前記デバイスおよび前記制御デバイスに接続された制御装置を備え、該制御装置は、多重化されたループ状インタフェースの全ループ異常時、自制御装置が接続されたループ状インタフェースのループ異常が一定時間内に解消するか否かを判定し、一定時間内にループ異常が解消した場合には前記制御デバイスに全ての前記デバイスの切り離しを行ったか否かを問い合わせ、前記制御デバイスによって全ての前記デバイスの切り離しが行われていたときに前記ループ診断手段によるループ診断を含む両系リンク障害対策処理を実行するものであって良い。また、前記ループ診断手段は、前記ループ診断で故障と判断された前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させるものであって良い。また、前記ループ診断手段は、利用を再開した前記ループ状インタフェースに接続された前記制御デバイスを經由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するものであって良い。

【0015】

本発明のエンクロージャサービスデバイスは、多重化されたFC-A L等のループ状インタフェースの1つのループ状インタフェースに接続され、接続された前記ループ状インタフェースへのハードディスク装置等のデバイスの接続および切り離しを行うループ接続切替手段を制御する機能を有するエンクロージャサービスデバイスにおいて、他のループ状インタフェースに接続された他のエンクロージャサービスデバイスと相互に通信するためのインタフェースを備え、且つ、全

ての前記ループ状インタフェースが異常となったことを検出したときに前記ループ接続切替手段を制御して前記ループ状インタフェースに接続された全ての前記デバイスを切り離す処理を行う手段を備えたことを特徴とする。

【0016】

ここで、本発明のエンクロージャサービスデバイスは、自デバイスが接続されるループ状インタフェースを通じて定期的に送られてくるコマンドが途絶えたことを検出して前記インタフェースを通じて他の前記エンクロージャサービスデバイスに通知する手段を備え、全ての前記エンクロージャサービスデバイスにおいて前記コマンドが途絶えたことが検出されたときに全ての前記ループ状インタフェースが異常となったことを検出するものであって良い。

【0017】

本発明の制御装置は、多重化されたFC-A L等のループ状インタフェースの1つのループ状インタフェースを通じて、1以上のハードディスク装置等のデバイスと、前記デバイスのループ状インタフェースへの接続および切り離しを行うループ接続切替手段を制御する機能を有するエンクロージャサービスデバイス等の制御デバイスとに接続されたディスク制御装置等の制御装置において、多重化されたループ状インタフェースの全てが異常となったことを検出した前記制御デバイスによって少なくとも1つの前記ループ状インタフェースに接続された全ての前記デバイスが切り離されたという事象を確認する手段と、接続された全ての前記デバイスが切り離されてループ異常が解消した前記ループ状インタフェースに接続された前記制御デバイスを経由して、前記制御デバイス間を相互に通信可能に接続するインタフェースを通じて、他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するループ診断手段を備えたことを特徴とする。

【0018】

ここで、本発明の制御装置における前記ループ診断手段は、前記ループ診断で故障と判断された前記デバイスを前記ループ状インタフェースから切り離して当該ループ状インタフェースの利用を再開させるものであって良い。また、前記ループ診断手段は、利用を再開した前記ループ状インタフェースに接続された前記制

御デバイスを経由して他の前記ループ状インタフェースに接続された前記制御デバイスをアクセスして故障デバイスを特定するループ診断を実施するものであって良い。

【0019】

【作用】

本発明にあっては、多重化された全てのループ状インタフェースが異常になったことが、各ループ状インタフェースに接続された制御デバイスにおいて検出される。そして、その検出時に制御デバイスが自律的に、ループ接続切替手段を制御して少なくとも1つのループ状インタフェースに接続された全てのデバイスをループから切り離す。これにより、ループ障害の原因が何れかのデバイスの故障である場合、全てのデバイスがループから切り離されたループ状インタフェースのループ異常は解消し、そのループ状インタフェースに接続された制御デバイスへのアクセスが可能となる。制御デバイス間はインタフェースにより相互に通信可能になっているので、アクセス可能となった制御デバイスを経由することにより、他の制御デバイスもアクセス可能となる。このため、任意のループ状インタフェースに接続されたデバイスの切り離し、再接続が行える。これにより、ループ異常の原因となる故障デバイスを特定するループ診断を実施でき、故障デバイスを切り離すことによって、システムの運用を継続することができる。

【0020】

【発明の実施の形態】

次に本発明の実施の形態について図面を参照して詳細に説明する。

【0021】

【構成】

図1を参照すると、本発明の一実施の形態にかかるディスクアレイ装置1は、ホストインタフェースA121、B122を介してホストA91、B92と接続され、ディスクユニット部5と、キャッシュ部6と、ディスク制御部A71と、ディスク制御部B72と、ホスト制御部A81と、ホスト制御部B82とを含んで構成される。

【0022】

ホスト制御部A 8 1、B 8 2は、ホストA 9 1、B 9 2間とのコマンド受信、データ転送、ステータス応答などのホストサービスを行う。ホスト制御部A 8 1、B 8 2は、内部バス1 1 0によりキャッシュ部6およびディスク制御部A 7 1、B 7 2に接続されており、お互いにデータの送受を行う。またホスト制御部A 8 1、B 8 2とディスク制御部A 7 1、B 7 2は、制御部間通信手段1 0 0により、ディスク処理などの他の制御部に対する処理要求や、ループ診断結果の通知などの制御部の状態通知などを行う。

【0 0 2 3】

ディスクユニット部5は、FC-A Lインタフェース・ポートを2ポート有する複数のFC-A Lディスク2 1～2 Nと、ループ接続切替部A 3 1、B 3 2と、エンクロージャサービス部A 5 1、B 5 2とを含んで構成される。

【0 0 2 4】

ループ接続切替部A 3 1、B 3 2は、FC-A Lディスク2 1～2 NのFC-A L A 4 1、B 4 2からの切り離し（バイパス）や、FC-A L A 4 1、B 4 2への接続（バイパス解除）を行う回路である。

【0 0 2 5】

エンクロージャサービス部A 5 1、B 5 2は、ディスク制御部A 7 1、B 7 2に対してFC-A L A 4 1、B 4 2を通じて接続されている。また、エンクロージャサービス部A 5 1、B 5 2は、相互に通信するためのインタフェース1 3 0を有している。このエンクロージャサービス部A 5 1、B 5 2は、ANSI規格（NCITS 305-199X）で規定されている装置の筐体内環境監視・資源管理のためのデバイスであり、ハートビート監視部A 5 1 1、B 5 2 1と、ハートビート受信部A 5 1 2、B 5 2 2と、ループ接続制御部A 5 1 3、B 5 2 3とを有している。

【0 0 2 6】

ループ接続制御部A 5 1 3、B 5 2 3は、ループ接続切替部A 3 1、B 3 2に対して、FC-A Lディスク2 1～2 Nのバイパス、バイパス解除を制御する機能を有する。

【0 0 2 7】

ハートビート受信部 A 5 1 2、B 5 2 2 は、F C - A L A 4 1、B 4 2 を通じてディスク制御部 A 7 1、B 7 2 から一定時間間隔で送られてくるコマンドを受信する機能を有する。ディスク制御部 A 7 1、B 7 2 から一定時間間隔で送られてくるコマンドを、以下、ハートビートコマンドと呼ぶ。ハートビートコマンドはそれ専用のコマンドを定義しても良いし、既存のコマンド（例えば R e c e i v e D i a g n o s t i c R e s u l t s コマンドなど）であっても良い。

【0028】

ハートビート監視部 A 5 1 1、B 5 2 1 は、自エンクロージャサービス部のハートビート受信部 A 5 1 2、B 5 2 2 におけるハートビートコマンドの受信状態を監視し、ハートビートコマンドの受信が途絶えたことを検出すると、インタフェース 1 3 0 を通じて他方のハートビート監視部 B 5 2 1、A 5 1 1 へその旨を通知する。そして、両ハートビート受信部 A 5 1 2、B 5 2 2 の双方でハートビートコマンドの受信が途絶えたことを検出すると、ループ接続制御部 A 5 1 3、B 5 2 3 によって、全ての F C - A L ディスク 2 1 ~ 2 N を F C - A L A 4 1、B 4 2 からバイパスする機能を有する。また、エンクロージャサービス部 A 5 1、B 5 2 は、コマンド受信が途絶えたことにより全ディスクのバイパスを実施したときには、その旨を内部に記録し、ディスク制御部 A 7 1、B 7 2 からの問い合わせに対して、バイパスを実施したことを報告する機能を有する。この問い合わせに対する報告機能は、例えば、R e c e i v e D i a g n o s t i c R e s u l t s コマンドに対する応答機能として実現することができる。

【0029】

エンクロージャサービス部 A 5 1 の内部構成例を図 2 に示す。エンクロージャサービス部 B 5 2 も同様の構成を有する。この例では、エンクロージャサービス部 A 5 1 は、C P U 5 3 1 と、そのバス 5 3 2 に接続されたインタフェースチップ 5 3 3、メモリ 5 3 4、通信部 5 3 5 および切替部 5 3 6 とを含んで構成される。インタフェースチップ 5 3 3 は、F C - A L A 4 1 との間のインタフェース部分である。通信部 5 3 5 はインタフェース 1 3 0 を通じてエンクロージャサービス部 B 5 2 と相互に通信するためのものである。切替部 5 3 6 は、ループ接続切替部 A 3 1 に対する制御信号等を保持するレジスタ等で構成される。メモリ 5

34は、ROMおよびRAMで構成され、エンクロージャサービスデバイス用プログラム（ファームウェアを含む）を記憶すると共に、自デバイスの状態やFC-A L A41に接続された各ディスクの状態などを示す状態情報等を記憶するための領域を有する。CPU531は、メモリ534に記憶されたエンクロージャサービスデバイス用プログラムを実行することにより、本デバイス全体を制御し、本デバイスに必要な機能を実現する。

【0030】

再び図1を参照すると、ディスク制御部A71は、FC-A L ディスク21～2Nの一方のポートのみとFC-A L A41を構成し、別のディスク制御部B72は、FC-A L ディスク21～2Nのもう一方のポートのみとFC-A L B42を構成する。ディスク制御部A71、B72は、ホスト制御部A81、B82からの指示により、あるいはディスク制御部A71、B72自身の判断により、FC-A L ディスク21～2Nに対してリードやライト等の各種ディスク処理を行う。また、ディスク制御部A71、B72は、エンクロージャサービス部A51、B52に対してFC-A L ディスク21～2Nのバイパス、バイパス解除の指示を出したり、バイパス状態などの情報をエンクロージャサービス部A51、B52から読み出したりする。更に、ディスク制御部A71、B72は、仮縮退制御部A710、B720と、ループ診断部A711、B721と、ハートビート送信部A712、B722とを有している。

【0031】

ハートビート送信部A712、B722は、エンクロージャサービス部A51、B52の構成要素であるハートビート受信部A512、B522に対して、一定時間間隔ごとにハートビートコマンドを発行する。前述したようにエンクロージャサービス部A51、B52の構成要素であるハートビート受信部A512、B522は発行されたハートビートコマンドを受信し、ハートビート監視部A511、B521は、ハートビートコマンドの受信状態を監視し、両ハートビート受信部A512、B522ともハートビートコマンドが途絶えたことを検出すると、全FC-A L ディスク21～2Nが双方のFC-A L A41、B42から切り離される。ハートビート監視部A511、B522がハートビートコマンド受

信が途絶えたと判断する時間は、ディスク制御部 A 7 1、B 7 2 から設定可能である。ハートビート送信部 A 7 1 2、B 7 2 2 は、ハートビート監視部 A 5 1 1、B 5 2 2 によるコマンド受信停止判断時間より短い間隔でハートビート受信部 A 5 1 2、B 5 2 2 に対してコマンドを発行する。

【0032】

ループ診断部 A 7 1 1、B 7 2 1 は、ループ異常が発生した場合に、F C - A L ディスクのバイパス、バイパス解除を伴うループ診断を互いに連携して実行し、故障ディスクを特定する機能を有する。

【0033】

仮縮退制御部 A 7 1 0、B 7 2 0 は、ループ診断中もホスト A 9 1、B 9 2 に対する応答を継続するために、ホスト制御部 A 8 1、B 8 2 やディスク制御部 A 7 1、B 7 2 間の連携を制御する機能を有する。

【0034】

ディスク制御部 A 7 1 の内部構成例を図 3 に示す。ディスク制御部 B 7 2 も同様の構成を有する。この例では、ディスク制御部 A 7 1 は、C P U 7 3 1 と、そのバス 7 3 2 に接続されたインタフェースチップ 7 3 3、メモリ 7 3 4、通信部 7 3 5 および DMA (ダイナミックメモリアクセス) コントローラ等で構成されるデータ転送部 7 3 6 とを含んで構成される。インタフェースチップ 7 3 3 は、F C - A L A 4 1 との間のインタフェース部分である。ディスク制御部におけるループ異常検出機構は、一般にこのインターフェースチップ 7 3 3 に設けられている。通信部 7 3 5 は制御部間通信手段 1 0 0 を通じて他の制御部と相互に通信するためのものである。データ転送部 7 3 6 は、内部バス 1 1 0 を通じて他の制御部とデータの授受を行うと共にインタフェースチップ 7 3 3 および F C - A L

A 4 1 を通じてエンクロージャサービス部 5 1 A や各ディスク 2 1 ~ 2 N とデータの授受を行う部分である。メモリ 7 3 4 は、R O M および R A M で構成され、ディスク制御部用プログラム (ファームウェアを含む) 等を記憶する。C P U 7 3 1 は、メモリ 7 3 4 に記憶されたディスク制御部用プログラムを実行することにより、本ディスク制御部全体を制御し、本ディスク制御部に必要な機能を実現する。

【0035】

【動作】

次に本実施の形態のディスクアレイ装置の動作について説明する。

【0036】

図1を参照すると、ホストA91、B92から命令を受け取ったホスト制御部A81、B82は、論理ディスク番号（LUN）、命令コード種別、および論理ブロックアドレス（LBA）などの必要な情報を認識する。例えば、ホストA91、B92からのリード命令を受けたホスト制御部A81、B82は、キャッシュ部6に指定されたデータが存在すれば、直ちにキャッシュ部6からホストA91、B92にデータを転送する。キャッシュ部6にデータが存在しなければ、FC-A Lディスク21～2Nから読み出したデータをキャッシュ部6に格納するようディスク制御部A71、B72に対して指示し、キャッシュ部6に格納されたら、そのデータをホストA91、B92に転送する。また例えば、ホストA91、B92からのライト命令を受けたホスト制御部A81、B82は、ホストA91、B92から受け取ったデータをキャッシュ部6に格納する。このデータはホスト制御部A81、B82からディスク制御部A71、B72に対してディスクへの書き込み指示が出されることにより、あるいはディスク制御部A71、B72がディスクへの未書き込みデータがキャッシュ部6に存在することを検出することにより、ディスク制御部A71、B72によってFC-A Lディスク21～2Nへ書き込まれる。2つのディスク制御部A71、B72は、処理を担当するFC-A Lディスク21～2Nを分担して負荷を分散させる使い方が一般的である。

【0037】

次に、FC-A L A41、FC-A L B42がループ異常を起こした場合の処理を説明する。

【0038】

FC-A L A41にのみループ異常が発生すると、FC-A L A41上を信号が伝搬できなくなるため、従来と同様にディスク制御部A71がそのことを検出する。このとき、ディスク制御部A71のハートビート送信部A712が定期

的に送信するハートビートコマンドもエンクロージャサービス部 A 5 1 のハートビート受信部 A 5 1 2 で受信されなくなり、ハートビート監視部 A 5 1 1 は FC-A L A 4 1 にループ異常が発生したことを検知するが、他方の FC-A L B 4 2 は正常なので、FC-A L A 4 1 からディスク 2 1 ~ 2 N をバイパスする制御は実施されない。

【0039】

反対に、FC-A L B 4 2 にのみループ異常が発生すると、FC-A L B 4 2 上を信号が伝搬できなくなるため、従来と同様にディスク制御部 B 7 2 がそのことを検出する。このとき、ディスク制御部 B 7 2 のハートビート送信部 B 7 2 2 が定期的に送信するハートビートコマンドもエンクロージャサービス部 B 5 2 のハートビート受信部 B 5 2 2 で受信されなくなり、ハートビート監視部 B 5 2 1 は FC-A L B 4 2 にループ異常が発生したことを検知するが、他方の FC-A L A 4 1 は正常なので、FC-A L B 4 2 からディスク 2 1 ~ 2 N をバイパスする制御は実施されない。

【0040】

FC-A L A 4 1、B 4 2 の双方でループ異常が発生すると、ディスク制御部 A 7 1 が FC-A L A 4 1 のループ異常を検出し、ディスク制御部 B 7 2 が FC-A L B 4 2 のループ異常を検出する。また、エンクロージャサービス部 A 5 1、B 5 2 のハートビート監視部 A 5 1 1、B 5 2 1 が、FC-A L A 4 1、B 4 2 の双方にループ異常が発生したことを検知し、ループ接続制御部 A 5 1 3 からループ接続切替部 A 3 1 が制御されて FC-A L A 4 1 から全てのディスク 2 1 ~ 2 N がバイパスされ、同時にループ接続制御部 B 5 2 3 からループ接続切替部 B 3 2 が制御されて FC-A L B 4 2 から全てのディスク 2 1 ~ 2 N がバイパスされる。

【0041】

図 4 はディスク制御部 A 7 1、B 7 2 が自ループの異常を検出したときの処理例を示すフローチャートである。ディスク制御部 A 7 1、B 7 2 と同じ処理を行うので、以下では、ディスク制御部 A 7 1 を例に説明する。

【0042】

ディスク制御部A71は、FC-A L A41のループ異常を検出すると(S101)、制御部間通信手段100を通じてディスク制御部B72に他方のFC-A L B42のループ状態を問い合わせる(S102)。ディスク制御部B72は、この問い合わせに対して、FC-A L B42のループ状態が正常か、異常かを示す問い合わせ応答を制御部間通信手段100を通じてディスク制御部A71に返却する。ディスク制御部A71は、他方のFC-A L B42がループ異常でなく正常であれば(S103でNO)、1つのループのみでループ異常が発生した場合の片系リンク障害対策処理を実施する(S104)。この片系リンク障害対策処理の詳細は後述する。他方、FC-A L B42もループ異常を起こしていれば(S103でYES)、ステップS105へ進む。

【0043】

ステップS105では、ディスク制御部A71は、一定時間だけ待ち合わせを行う。待ち合わせる時間は、エンクロージャサービス部A51、B52がFC-A L A41およびB42の双方でループ異常が発生した場合に自律的に全てのディスク21~2NをFC-A L A41、B42からバイパスさせるのに要する時間より若干長い時間である。この待ち合わせ後、ディスク制御部A71は、FC-A L A41のループ異常が解消しているかどうかを判別する(S106)。そして、自ループの異常が解消していれば(S106でYES)、ループ異常が解消したFC-A L A41を通じてエンクロージャサービス部A51に対して、ハートビートコマンドの受信が両ループで途絶えたために全ディスク21~2Nのバイパスを実施したかどうかを問い合わせる(S107)。この問い合わせに対して、全ディスクのバイパスを実施した旨の回答がエンクロージャサービス部A51からFC-A L A41を通じて返された場合(S108でYES)、FC-A L A41とB42の両方でループ異常が固定的に発生した場合の両系リンク障害対策処理を実施する(S109)。この両系リンク障害対策処理の詳細は後述する。

【0044】

他方、エンクロージャサービス部A51から全ディスクのバイパスを実施していない旨の回答が得られた場合(S108でNO)、発生したループ異常が間欠障

害で、自然に治癒したものとして処理する。また、ステップ S105 で一定時間待ち合わせても、FC-A L A41 のループ異常が解消しなかった場合 (S106 で NO)、ループ異常の原因が例えばエンクロージャサービス部 A51 自体の故障等、ディスク 21 ~ 2N 以外の箇所に存在するので、それに応じた障害対策処理を実施する。この障害対策処理は本発明と直接関係しないので、その説明は省略する。

【0045】

図4の処理例では、他ループも異常の場合 (S103 で YES)、一定時間だけ待ってから (S105)、自ループの異常が解消したかを判断したが (S106)、この部分の処理を図5のように変更しても良い。つまり、他ループも異常の場合 (S103 で YES)、自ループの異常が解消したかどうかを判断する処理 S106 と一定時間経過したかどうかを判断する処理 S110 とを含むループを実行し、自ループの異常が解消していればステップ S107 へ進み、一定時間経過しても自ループの異常が解消していなければ (S110 で YES)、他の障害対策処理へ進む。図4の処理では、ループ障害が間欠障害であって直ぐに解消した場合も一定時間待つ必要があるが、図5の処理によれば、そのようなことがなくなる利点がある。

【0046】

次に、ステップ S104 の片系リンク障害対策処理について、FC-A L A41 がループ異常となった場合を例に説明する。当然ながら、FC-A L B42 の場合も同様に動作する。

【0047】

図6は片系リンク障害対策処理 (S104) の一例を示すフローチャートである。例えば、何れかの FC-A L ディスク 21 ~ 2N の故障により FC-A L A41 にてリンクダウンなどのループ異常が継続したとする (S1041)。この状態ではディスク制御部 A71 は、エンクロージャサービス部 A51 や FC-A L ディスク 21 ~ 2N に対してアクセスができなくなる。ディスク制御部 A71 の仮縮退制御部 A710 は、通常のディスク処理の実行を停止 (この通常機能を停止している状態を仮縮退状態と呼ぶこととする) するとともに (S1042)

、制御部間通信手段100によりホスト制御部A81、B82および別のディスク制御部B72に対して仮縮退状態に移したことを通知（仮縮退通知）する（S1043）。

【0048】

仮縮退通知を受けたディスク制御部B72、ホスト制御部A81、B82は、ループA仮縮退状態でシステムを運用する（S1044）。具体的には、仮縮退通知を受けたディスク制御部B72の仮縮退制御部B720は、最初にFC-A Lディスク21～2Nのリセットを実行し、仮縮退したディスク制御部A71が放棄したディスク処理によりFC-A Lディスク21～2N内に滞留している処理をキャンセルする。またディスク制御部B72は、全てのFC-A Lディスク21～2Nを対象に、ホスト制御部A81、B82から指示されたディスク処理、およびディスク制御部B72の判断によるディスク処理を行う。他方、仮縮退通知を受けたホスト制御部A81、B82は、仮縮退したディスク制御部A71に対して要求していた未完了のディスク処理を代替のディスク制御部B72に対して再度要求する。またディスク制御部A71が仮縮退中は、新たなホストI/Oによるディスク処理について全て代替のディスク制御部B72に対して要求する。よって、ディスク制御部A71が仮縮退している間もディスク制御部B72が処理を引き継ぐため、ホストに対して応答を継続することが可能となる。

【0049】

次に、仮縮退したディスク制御部A71のループ診断部A711とディスク制御部B72のループ診断部B721とは、連携してFC-A L A41に接続された複数のディスク21～2Nから故障部品を特定するための診断処理を行う（S1045）。以下、ステップS1045のループ診断処理の具体例を説明する。

【0050】

まず、ディスク制御部A71のループ診断部A711は、制御部間通信手段100を通じてディスク制御部B72のループ診断部B721に対して、FC-A Lディスク21～2Nのうち、例えばディスク21をFC-A L A41からバイパスするよう指示する。指示を受けたディスク制御部B72は、FC-A L B42を通じて、エンクロージャサービス部B52に対して、同様のバイパスを指

示し、エンクロージャサービス部 B 5 2 は、インタフェース 1 3 0 を通じてエンクロージャサービス部 A 5 1 に同様のバイパスを指示する。エンクロージャサービス部 A 5 1 のループ接続制御部 A 5 1 3 は、この指示を受けてループ接続切替部 A 3 1 を制御し、ディスク 2 1 を FC-A L A 4 1 から切り離す。ディスク制御部 A 7 1 は、FC-A L A 4 1 のループ異常が解消したかどうかを判定し、解消していれば、ディスク 2 1 を故障ディスクと特定する。他方、ループ異常がなおも継続していれば、次のディスク 2 2 についてディスク 2 1 と同様の手順で FC-A L A 4 1 からバイパスし、ディスク 2 2 が故障ディスクでないかどうかを調べる。このような処理を故障ディスクが特定されるまで、最後のディスク 2 N まで繰り返す。

【0051】

こうして故障ディスクを特定すると、ディスク制御部 A 7 1 は、故障ディスクは切り離し、正常ディスクはバイパス解除して FC-A L A 4 1 に接続し直す（ステップ S 1 0 4 6）。例えば、ディスク 2 2 が故障ディスクであった場合、図 7 に示すように、FC-A L A 4 1 から故障ディスク 2 2 を切り離し、正常ディスクだけを接続する。

【0052】

その後、ディスク制御部 A 7 1 は、制御部間通信手段 1 0 0 によりホスト制御部 A 8 1、B 8 2 および別のディスク制御部 B 7 2 に対して、仮縮退状態を解除し通常状態に遷移したことを通知（仮縮退解除通知）する（S 1 0 4 7）。仮縮退状態を解除したディスク制御部 A 7 1 は、通常機能としてのディスク処理を再開する。仮縮退解除通知を受けたディスク制御部 B 7 2 は、仮縮退解除したディスク制御部 A 7 1 が受け持つディスクに対する処理を停止し、ループ診断により他方のループからバイパスされたディスクを含めた残りのディスクに対する処理を受け持つ。仮縮退解除通知を受けたホスト制御部 A 8 1、B 8 2 は、上記のディスクの受け持ちに応じたディスク制御部 A 7 1、B 7 2 に対してディスク処理を要求する。故障と判断されたディスクは、例えばディスクアレイ装置 1 に接続された保守端末（図示せず）や、あるいはホストバスで接続されたディスクアレイ管理用ホスト（図示せず）などにより保守員や使用者などに、交換が必要である

ことが通知される。

【0053】

上述したループ診断処理の具体例はあくまでも一例であり、本発明は上記具体例に限定されるものではない。例えば、文献2の図5に示されるような手順によるループ診断を行っても良い。また、FC-AL A41に接続されている全てのディスク21～2Nを一旦バイパスし、1つのディスク毎にバイパス解除してそのディスクが故障ディスクかどうかを調べるようにしても良い。

【0054】

次に、図4の両系リンク障害対策処理(S109)について説明する。

【0055】

図8は両系リンク障害対策処理S109の一例を示すフローチャートである。この両系リンク障害対策処理が開始される前には、エンクロージャサービス部A51、B52によってFC-AL A41、B42から全てのディスク21～2Nがバイパスされており、この状態ではディスク制御部A71、B72は共に通常のディスク処理が行えない。このため、ディスク制御部A71、B72の仮縮退制御部A710、B720は、通常のディスク処理の実行を停止して仮縮退状態とするとともに(S1091)、制御部間通信手段100によりホスト制御部A81、B82に対して仮縮退状態に遷移したことを通知(仮縮退通知)する(S1092)。仮縮退通知を受けたホスト制御部A81、B82は、全てのディスク制御部A71、B72が仮縮退状態となったので、ホストA91、B92からの要求の受付を一時的に中断する(S1093)。

【0056】

次に、仮縮退したディスク制御部A71、B72のループ診断部A711、B721は、先ず、何れか一方のFC-AL、例えばFC-AL A41に接続された複数のディスク21～2Nから故障部品を特定するための診断処理を連携して行う(S1094)。以下、ステップS1094のループ診断処理の具体例を説明する。

【0057】

先ず、ディスク制御部A71のループ診断部A711は、制御部間通信手段10

0を通じてディスク制御部B72のループ診断部B721に対して、FC-A L ディスク21～2Nのうち、例えばディスク21をFC-A L A41に接続（バイパス解除）するよう指示する。指示を受けたディスク制御部B72は、FC-A L B42を通じて、エンクロージャサービス部B52に対して、同様のバイパス解除を指示し、エンクロージャサービス部B52は、インタフェース130を通じてエンクロージャサービス部A51に同様のバイパス解除を指示する。エンクロージャサービス部A51のループ接続制御部A513は、この指示を受けてループ接続切替部A31を制御し、ディスク21をFC-A L A41に接続する。ディスク制御部A71は、FC-A L A41のループ異常が再発したかどうかを判定し、再発していれば、ディスク21を故障ディスクと特定する。他方、ループ異常が再発しなければ、ディスク21は正常ディスクと特定する。

【0058】

ディスク21を故障ディスクと特定した場合、ディスク制御部A71のループ診断部A711は、そのバイパス解除を指示した場合と同じ経路、つまり、制御部間通信手段100、ループ診断部B721、エンクロージャサービス部B52、インタフェース130を経由する経路で、エンクロージャサービス部A51にディスク21をバイパスする指示を出し、故障ディスク21をFC-A L A41から切り離した後、次のディスク22の診断をディスク21と同様に実施する。ディスク21が故障ディスクでない場合は、ディスク21をFC-A L A41に接続したまま、次のディスク22の診断をディスク21と同様に実施する。このような手順を、残りの全てのディスクについて繰り返す。こうしてFC-A L A41のループ診断が終了すると、FC-A L A41には正常ディスクだけが接続された状態になる。

【0059】

なお、以上の説明では、ディスク制御部A71がディスク21～2NをFC-A L A41に接続する場合、他方のディスク制御部B72、FC-A L B42、エンクロージャサービス部B52経由でエンクロージャサービス部A51にバイパス解除の指示を伝達したが、FC-A L A41経由で直接にエンクロージャサービス部A51にバイパス解除の指示を送るようにしても良い。但し、若し、

バイパス解除したディスクが故障ディスクであれば、FC-A L A 4 1 は再びループ異常になるので、そのディスクを再び切り離す際には、他方のディスク制御部 B 7 2、FC-A L B 4 2、エンクロージャサービス部 B 5 2 経由でエンクロージャサービス部 A 5 1 にバイパスの指示を伝達する必要がある。

【0060】

その後、ディスク制御部 A 7 1 は、制御部間通信手段 100 によりホスト制御部 A 8 1、B 8 2 に対して、ループ A の仮縮退状態を解除し通常状態に遷移したことを通知（仮縮退解除通知）する（S 1095）。これにより、ループ B 仮縮退状態でシステムの運用が再開される（S 1096）。つまり、ループ A の仮縮退解除通知を受けたホスト制御部 A 8 1、B 8 2 は、仮縮退したディスク制御部 A 7 1、B 7 2 に対して要求していた未完了のディスク処理を、仮縮退解除したディスク制御部 A 7 1 に対して再度要求する。また、ホスト A 9 1、B 9 2 からの要求の受け付けを再開し、新たなホスト I/O によるディスク処理について全てディスク制御部 A 7 1 に対して要求する。ディスク制御部 A 7 1 は、正常な全ての FC-A L ディスクを対象に、それらに滞留しているディスク処理をキャンセルした後、ホスト制御部 A 9 1、B 9 2 から指示されたディスク処理、およびディスク制御部 A 7 1 の判断によるディスク処理を行う。これにより、システムの運用が完全に停止する期間を最小限にすることができる。

【0061】

次に、仮縮退解除したディスク制御部 A 7 1 のループ診断部 A 7 1 1 と、仮縮退中のディスク制御部 B 7 2 のループ診断部 B 7 2 1 は、連携して FC-A L B 4 2 に接続された複数のディスク 2 1 ~ 2 N から故障部品を特定するための診断処理を行う（S 1097）。このステップ S 1097 のループ診断処理は、例えば、ステップ S 1094 の処理と同様に以下のように行われる。

【0062】

まず、ディスク制御部 B 7 2 のループ診断部 B 7 2 1 は、制御部間通信手段 100、ディスク制御部 A 7 1 のループ診断部 7 1 1、FC-A L A 4 1、エンクロージャサービス部 A 5 1、インタフェース 130 を通じて、エンクロージャサービス部 B 5 2 に対して、FC-A L ディスク 2 1 ~ 2 N のうち、例えばディス

ク 2 1 を F C - A L B 4 2 に接続 (バイパス解除) するよう指示する。エンクロージャサービス部 B 5 2 のループ接続制御部 B 5 2 3 は、この指示を受けてループ接続切替部 B 3 2 を制御し、ディスク 2 1 を F C - A L B 4 2 に接続する。ディスク制御部 B 7 2 は、F C - A L B 4 2 のループ異常が再発したかどうかを判定し、再発していれば、ディスク 2 1 を故障ディスクと特定し、ループ異常が再発しなければ、ディスク 2 1 は正常ディスクと特定する。ディスク 2 1 を故障ディスクと特定した場合、ディスク制御部 B 7 2 のループ診断部 B 7 2 1 は、そのバイパス解除を指示した場合と同じ経路で、ループ接続切替部 B 3 2 にディスク 2 1 をバイパスする指示を出し、故障ディスク 2 1 を F C - A L B 4 2 から切り離した後、次のディスク 2 2 の診断をディスク 2 1 と同様に実施する。ディスク 2 1 が故障ディスクでない場合は、ディスク 2 1 を F C - A L B 4 2 に接続したまま、次のディスク 2 2 の診断をディスク 2 1 と同様に実施する。このような手順を、残りの全てのディスクについて繰り返す。こうして F C - A L B 4 2 のループ診断が終了すると、F C - A L B 4 2 には正常ディスクだけが接続された状態になる。なお、ステップ S 1 0 9 4 の場合と同様に、ディスク制御部 B 7 2 は F C - A L B 4 2 経由で直接にエンクロージャサービス部 B 5 2 にバイパス解除の指示を送るようにしても良い。

【0063】

その後、ディスク制御部 B 7 2 は、制御部間通信手段 1 0 0 によりホスト制御部 A 8 1、B 8 2 および他のディスク制御部 A 7 1 に対して、ループ B の仮縮退状態を解除し通常状態に遷移したことを通知 (仮縮退解除通知) する (S 1 0 9 8)。これにより、ループ A、B の双方を使用してシステムの運用が継続される (S 1 0 9 9)。つまり、仮縮退状態を解除したディスク制御部 B 7 2 は、通常機能としてのディスク処理を再開する。仮縮退解除通知を受けたディスク制御部 A 7 1 は、仮縮退解除したディスク制御部 B 7 2 が受け持つディスクに対する処理を停止する。ループ B の仮縮退解除通知を受けたホスト制御部 A 8 1、B 8 2 は、上記のディスクの受け持ちに応じたディスク制御部 A 7 1、B 7 2 に対してディスク処理を要求する。なお、故障と判断されたディスクは、例えばディスクアレイ装置 1 に接続された保守端末 (図示せず) や、あるいはホストバスで接続さ

れたディスクアレイ管理用ホスト（図示せず）などにより保守員や使用者などに、交換が必要であることが通知される。

【0064】

ステップS1094、S1097で説明した上述したループ診断処理の具体例はあくまでも一例であり、本発明は上記具体例に限定されるものではない。例えば、ステップS1094における診断方法として、例えば半分のディスクを一度にFC-A L A41に接続し、ループ異常が発生していなければ、残りのディスクの半分の一度にFC-A L A41に接続するというように、一度に複数のディスクを接続するようにしても良い。この場合、複数のディスクを接続した際にループ異常が再発したら、その中に故障ディスクが存在するので、それら複数のディスクの中から故障ディスクを特定する処理を行えば良い。ステップS1097においても同様のループ診断方法を用いることができる。

【0065】

また、実際にはディスク1台の故障で両ループが異常になるケースが多いので、ステップS1097では、ステップS1094の診断結果を考慮し、故障ディスクと判定されたディスク以外のディスク全てを一度にFC-A L B42に接続し、ループ異常が再発しないかどうかを調べるようにしても良い。但し、希ではあるが2台のディスクが1ループずつ同時に故障して2ループ異常が発生する場合もあり、その場合にはループ異常が再発するので、更に診断を続けてFC-A L B42側のループ異常の原因である故障ディスクを特定する。

【0066】

【発明の他の実施の形態】

以上の実施の形態では、エンクロージャサービス部A51、B52のハートビート監視部A511、B521が双方のループ異常を検出したとき、ループ接続制御部A513、B523によって、FC-A L A41から全ディスク21～2Nを切り離すと共に、FC-A L B42から全ディスク21～2Nを切り離すことで、双方のループの異常を解消させた。しかし、本発明はこれに限定されず、何れか一方のFC-A L、例えばFC-A L B42から全ディスク21～2Nを切り離し、他方のFC-A L A41にはディスクを接続したままにしてお

いても良い。この場合、FC-A L B 4 2 のループ異常は解消するが、FC-A L A 4 1 のループ異常は解消しない。この状態は、1つのループでのみ異常が発生した場合と類似した状態であり、図6のステップS 1 0 4 5で説明したと同様な方法で、ループ異常のないFC-A L B 4 2 経由でエンクロージャサーバ部5 1をディスク制御部A 7 1からアクセスすることで、FC-A L A 4 1に接続されたディスクの中から故障ディスクを特定するループ診断を実施できる。そして、ループ診断で故障と判断された故障ディスクをFC-A L A 4 1から切り離せば、FC-A L A 4 1のループ異常が解消するので、ステップS 1 0 4 5と同様な方法で、他方のFC-A L B 4 2のループ診断を実施すれば良い。

【0067】

以上の実施の形態では、ディスク制御部A 7 1、B 7 2に1つのFC-A Lが接続されたディスクアレイ装置に本発明を適用したが、本発明はこれに限定されず、ディスク制御部A 7 1、B 7 2に複数のFC-A Lが接続されたディスクアレイ装置に対しても本発明は適用可能である。図9にその一例を示す。

【0068】

図9に示すディスクアレイ装置は、ディスクユニット部5とそれと同様なディスクユニット部5 Xとの2つのディスクユニット部を備え、双方のディスクユニット部のFC-A L A 4 1がディスク制御部A 7 1に接続され、双方のディスクユニット部のFC-A L B 4 2がディスク制御部B 7 2に接続されている。つまり、2個の独立したディスクユニット部5、5 Xが並列にディスク制御部A 7 1、B 7 2に接続されている。ディスクユニット部5だけを備える図1のディスクアレイ装置では、ディスクユニット部5内のディスクを使ってディスクアレイを構成するが、図9のディスクアレイ装置では、それだけでなく、例えばディスクユニット部5のディスク2 1とディスクユニット部5 Xのディスク2 1とでRAID 1のディスクアレイを構成する如く、異なるディスクユニット部のディスクを組み合わせることでディスクアレイを構成することができる。この場合、何れか一方のディスクユニット部で2重ループ異常が発生しても、もう一方のディスクユニット部でディスク処理を継続することができるため、信頼性をより高めること

ができる。なお、各ディスクユニット部 5、5 X において 2 重ループ異常が発生した場合の処理は基本的に図 1 のディスクアレイ装置と同じである。

【0069】

以上の実施の形態では、1 つの FC-A L にエンクロージャサービス部が 1 つだけ接続されたディスクアレイ装置に本発明を適用したが、本発明はこれに限定されず、1 つの FC-A L に複数のエンクロージャサービス部が接続されたディスクアレイ装置に対しても本発明は適用可能である。その一例を図 10 に示す。

【0070】

図 10 に示すディスクアレイ装置は、ディスクユニット部 5 と同様なディスクユニット部 5 Y (増設装置) によって 1 つの FC-A L A 4 1、B 4 2 に接続するディスクの数を拡張している。この場合、ディスクユニット部 5 を含む基本装置とディスクユニット部 5 Y を含む増設装置毎にそれぞれの 1 ループ当たり 1 つのエンクロージャサービス部が設けられるので、ディスクユニット部 5 に存在する各ループのエンクロージャサービス部 5 1、5 2 同士、ディスクユニット部 5 Y 毎に存在する各ループのエンクロージャサービス部 5 1、5 2 同士をインタフェース 130 で互いに通信可能に接続する。また、ディスク制御部 A 7 1、B 7 2 からのハートビートコマンドは、同じループに接続された全てのエンクロージャサービス部に対して送信され、双方のループ異常を検出したときに各エンクロージャサービス部は自身が設けられている基本装置あるいは増設装置内のループ部分に接続されたディスク全てを切り離す。

【0071】

以上本発明の実施の形態について説明したが、本発明は以上の実施の形態にのみ限定されず、その他各種の付加変更が可能である。例えば、エンクロージャサービス部はハートビートコマンドの受信が途絶えたことでループ異常を検出したが、ディスク制御部におけるループ異常検出機能と同様な機能をエンクロージャサービス部に設けるようにしても良い。

【0072】

【発明の効果】

以上説明したように本発明によれば以下のような効果が得られる。

【0073】

ループへのデバイスの接続、切り離しを行うループ接続切替手段を制御するエンクロージャサービスデバイス等の制御デバイスがそのループに直接接続されたループ状インタフェースを多重化したシステムにおいて、全てのループ状インタフェースが同時にループ障害になった場合でも、デバイスの切り離し、再接続が行えることである。その理由は、多重化された全てのループ状インタフェースが異常になると、それが制御デバイスで検出され、制御デバイスが自律的にループ接続切替手段を制御して少なくとも1つのループ状インタフェースに接続された全てのデバイスをループから切り離し、少なくとも1つのループ状インタフェースのループ異常を解消して、そのループに接続された制御デバイスへのアクセス、この制御デバイス経由による他方のループに接続された制御デバイスへのアクセスが可能になるからである。

【0074】

ループへのデバイスの接続、切り離しを行うループ接続切替手段を制御するエンクロージャサービスデバイス等の制御デバイスがそのループに直接接続されたループ状インタフェースを多重化したシステムにおいて、全てのループ状インタフェースが同時にループ障害になった場合でも、ループ異常の原因となる故障デバイスを特定するループ診断が実施できることである。その理由は、ループ診断を実施するためには、ループ状インタフェースに接続されたデバイスの切り離し、再接続が必要であるが、このデバイスの切り離し、再接続が可能になるからである。

【0075】

ループへのデバイスの接続、切り離しを行うループ接続切替手段を制御するエンクロージャサービスデバイス等の制御デバイスがそのループに直接接続されたループ状インタフェースを多重化したシステムにおいて、全てのループ状インタフェースが同時にループ障害になった場合でも、システムが完全に停止したままになるのを防止できることである。その理由は、上述したようにループ診断が可能となり、故障部品をループから除去してシステムの運用を再開できるからである。

【図面の簡単な説明】**【図 1】**

本発明を適用したディスクアレイ装置の一例を示すブロック図である。

【図 2】

本発明を適用したディスクアレイ装置におけるエンクロージャサービス部の内部構成例を示すブロック図である。

【図 3】

本発明を適用したディスクアレイ装置におけるディスク制御部の内部構成例を示すブロック図である。

【図 4】

ディスクアレイ装置のディスク制御部がループの異常を検出したときの処理例を示すフローチャートである。

【図 5】

ディスクアレイ装置のディスク制御部がループの異常を検出したときの別の処理例を示すフローチャートである。

【図 6】

片系リンク障害対策処理の一例を示すフローチャートである。

【図 7】

故障ディスクを切り離し、正常ディスクだけを接続したディスクアレイ装置の状態を示す図である。

【図 8】

両系リンク障害対策処理の一例を示すフローチャートである。

【図 9】

本発明を適用したディスクアレイ装置の別の例を示すブロック図である。

【図 10】

本発明を適用したディスクアレイ装置の更に別の例を示すブロック図である。

【符号の説明】

1…ディスクアレイ装置

5、5 X、5 Y…ディスクユニット部

6…キャッシュ部
21～2N…FC-ALディスク
31…ループ接続切替部A
32…ループ接続切替部B
41…FC-AL A
42…FC-AL B
51…エンクロージャサービス部A
52…エンクロージャサービス部B
71…ディスク制御部A
72…ディスク制御部B
81…ホスト制御部A
82…ホスト制御部B
91…ホストA
92…ホストB
100…制御部間通信手段
110…内部バス
121…ホストインタフェースA
122…ホストインタフェースB
130…インタフェース
511…ハートビート監視部A
512…ハートビート受信部A
513…ループ接続制御部A
521…ハートビート監視部B
522…ハートビート受信部B
523…ループ接続制御部B
531…CPU
532…バス
533…インタフェースチップ
534…メモリ

5 3 5 …通信部

5 3 6 …切替部

7 1 0 …仮縮退制御部 A

7 1 1 …ループ診断部 A

7 1 2 …ハートビート送信部 A

7 2 0 …仮縮退制御部 B

7 2 1 …ループ診断部 B

7 2 2 …ハートビート送信部 B

7 3 1 …C P U

7 3 2 …バス

7 3 3 …インタフェースチップ

7 3 4 …メモリ

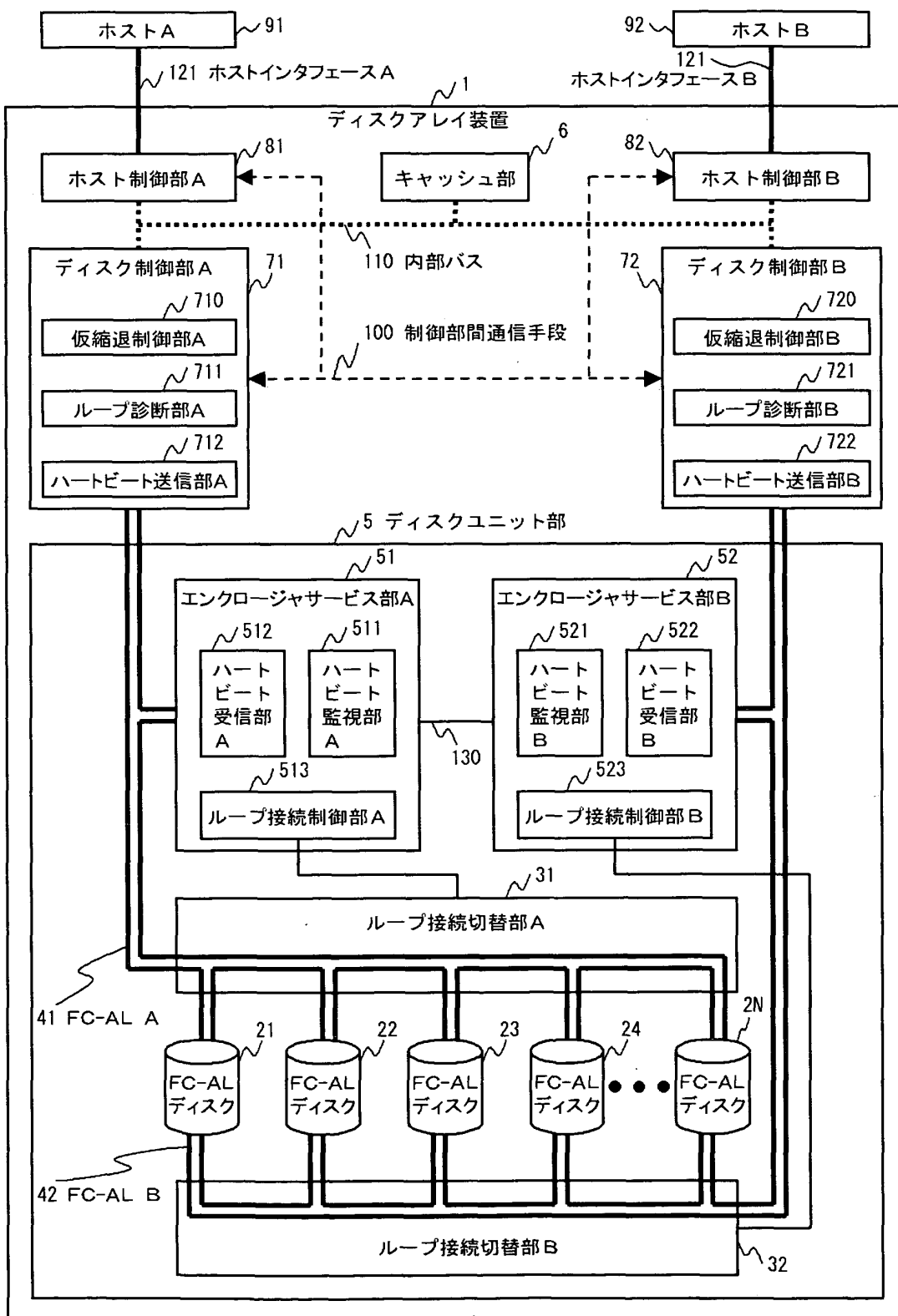
7 3 5 …通信部

7 3 6 …データ転送部

【書類名】 図面

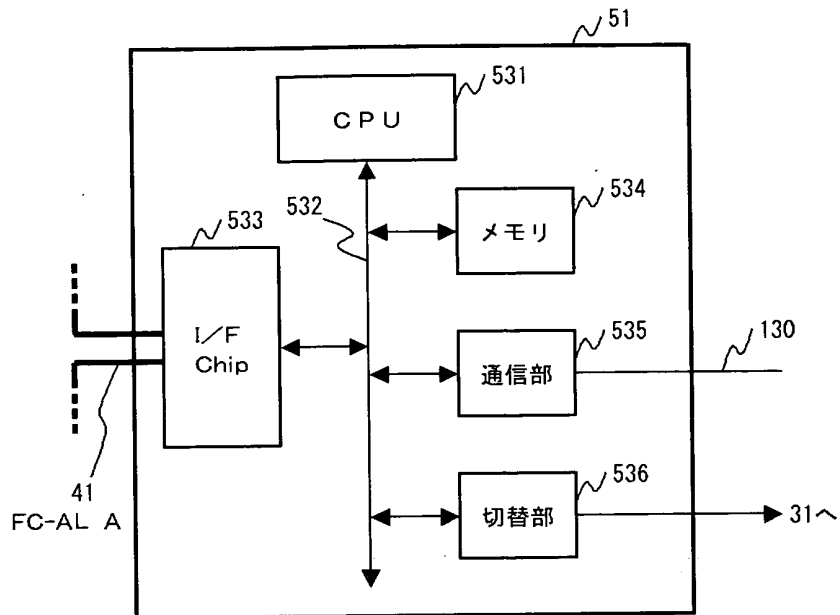
【図 1】

【図 1】



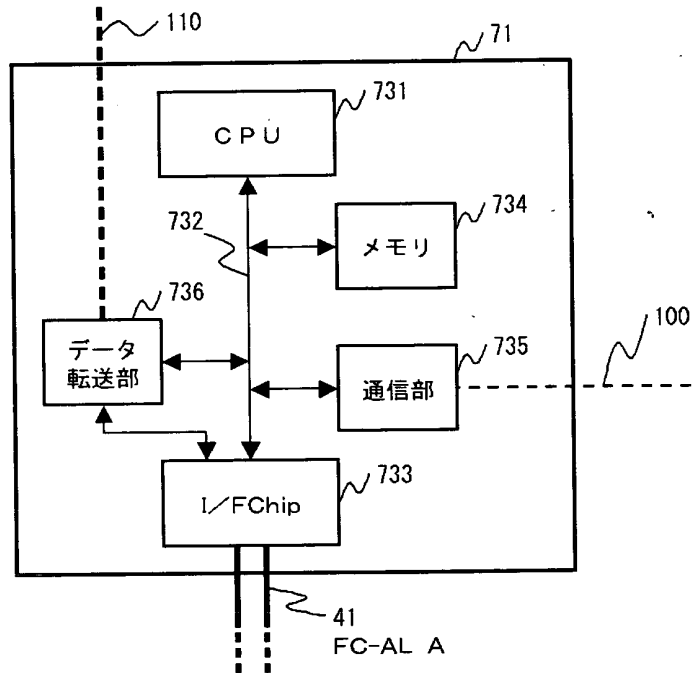
【図 2】

【図 2】



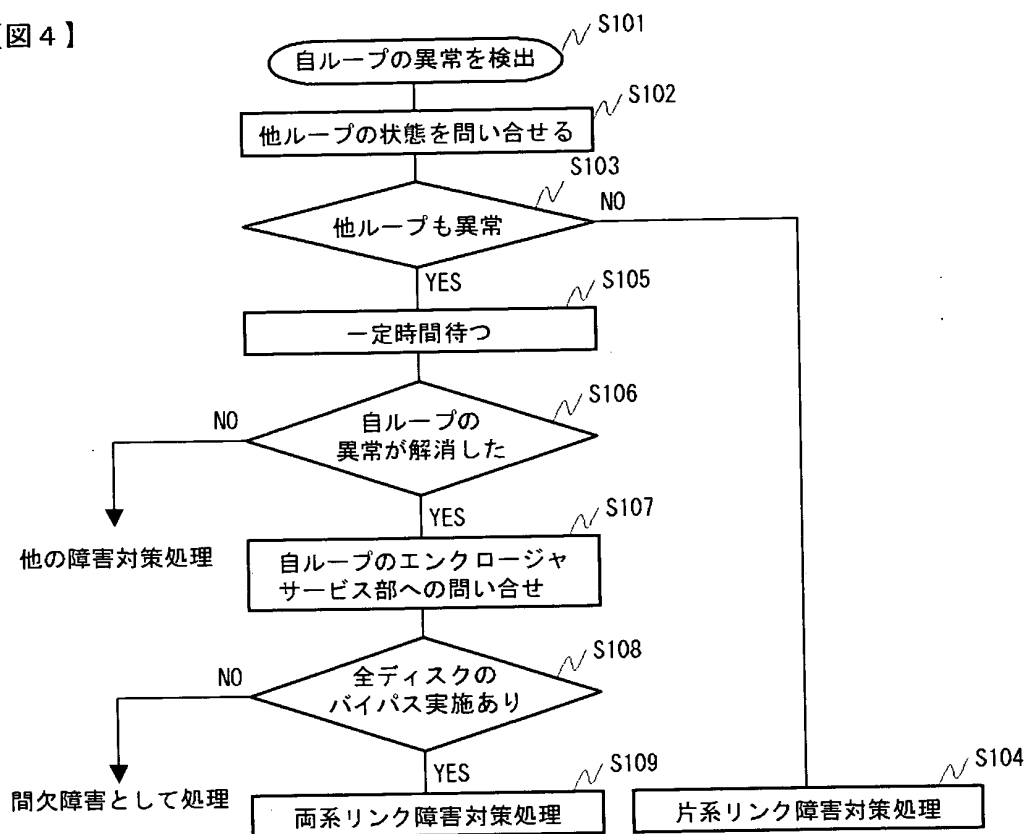
【図 3】

【図 3】



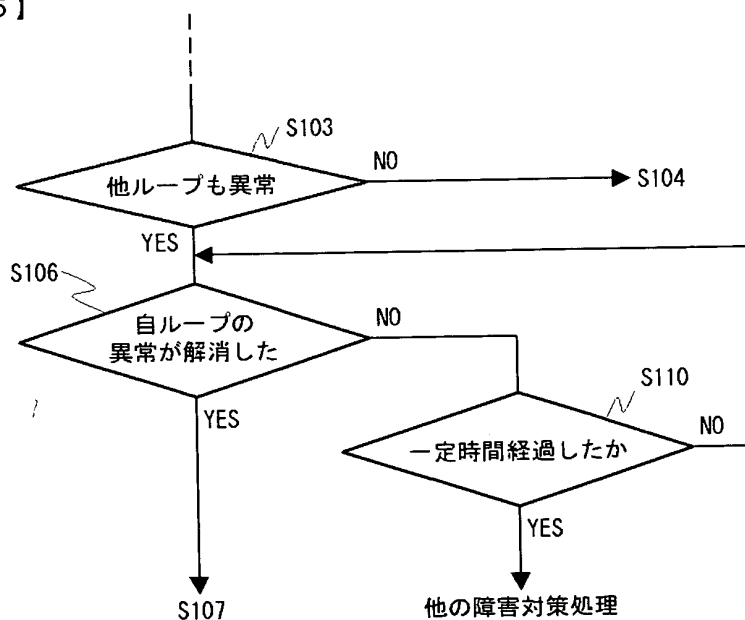
【図 4】

【図 4】



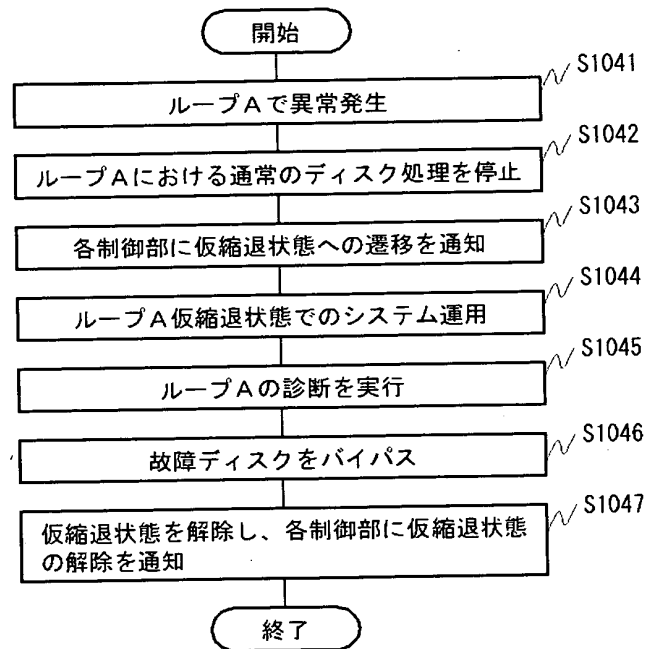
【図 5】

【図 5】



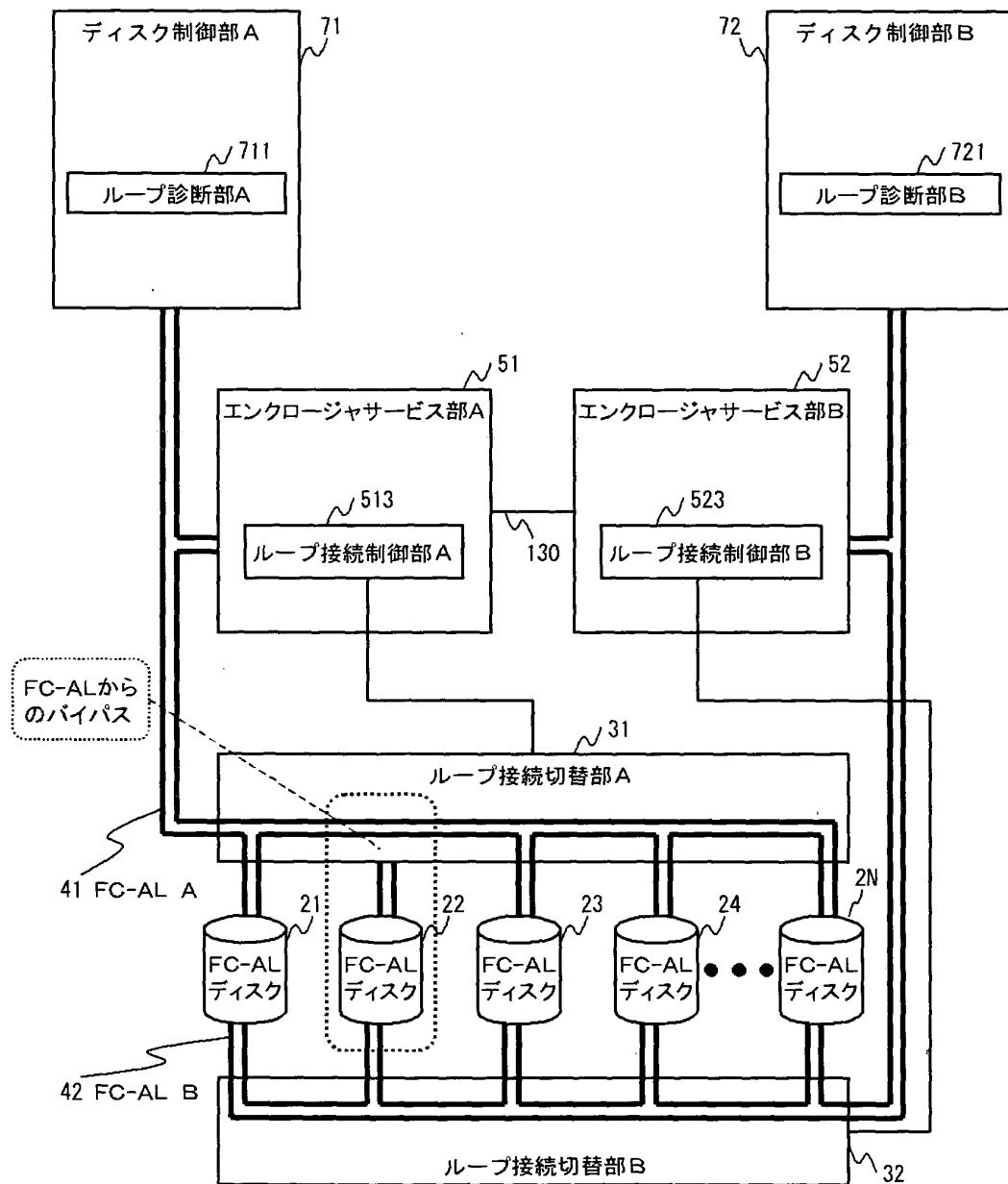
【図 6】

【図 6】



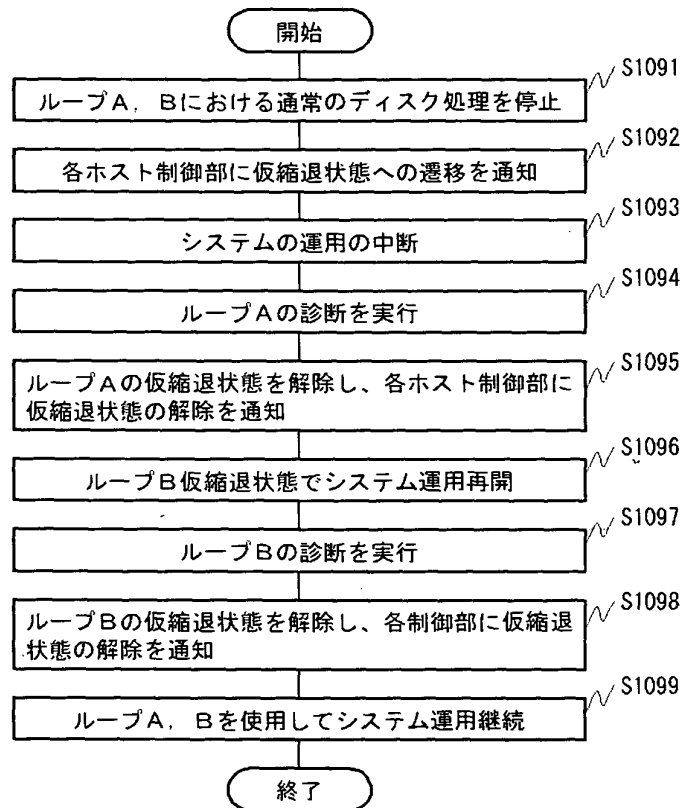
【図 7】

【図 7】



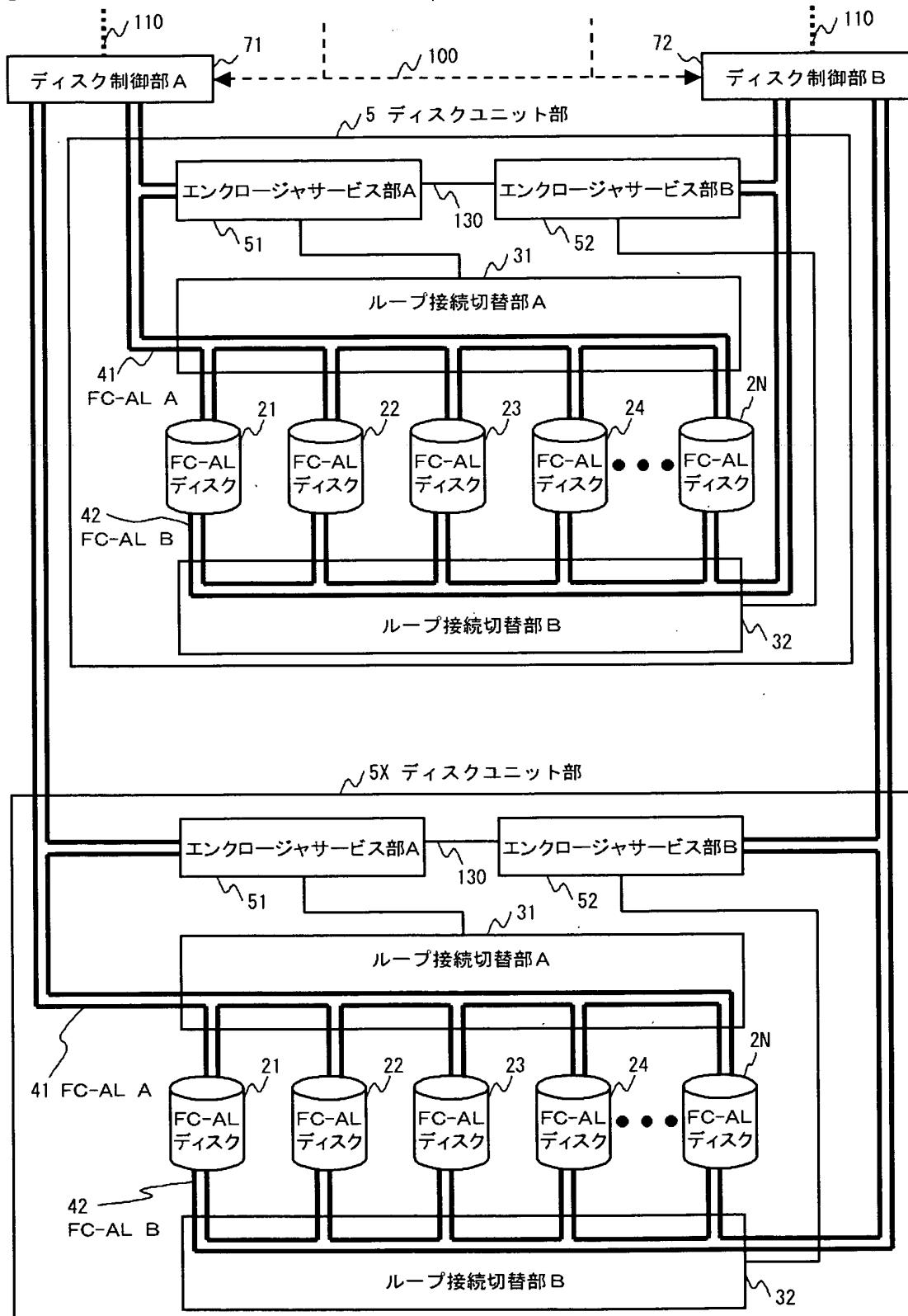
【図 8】

【図 8】



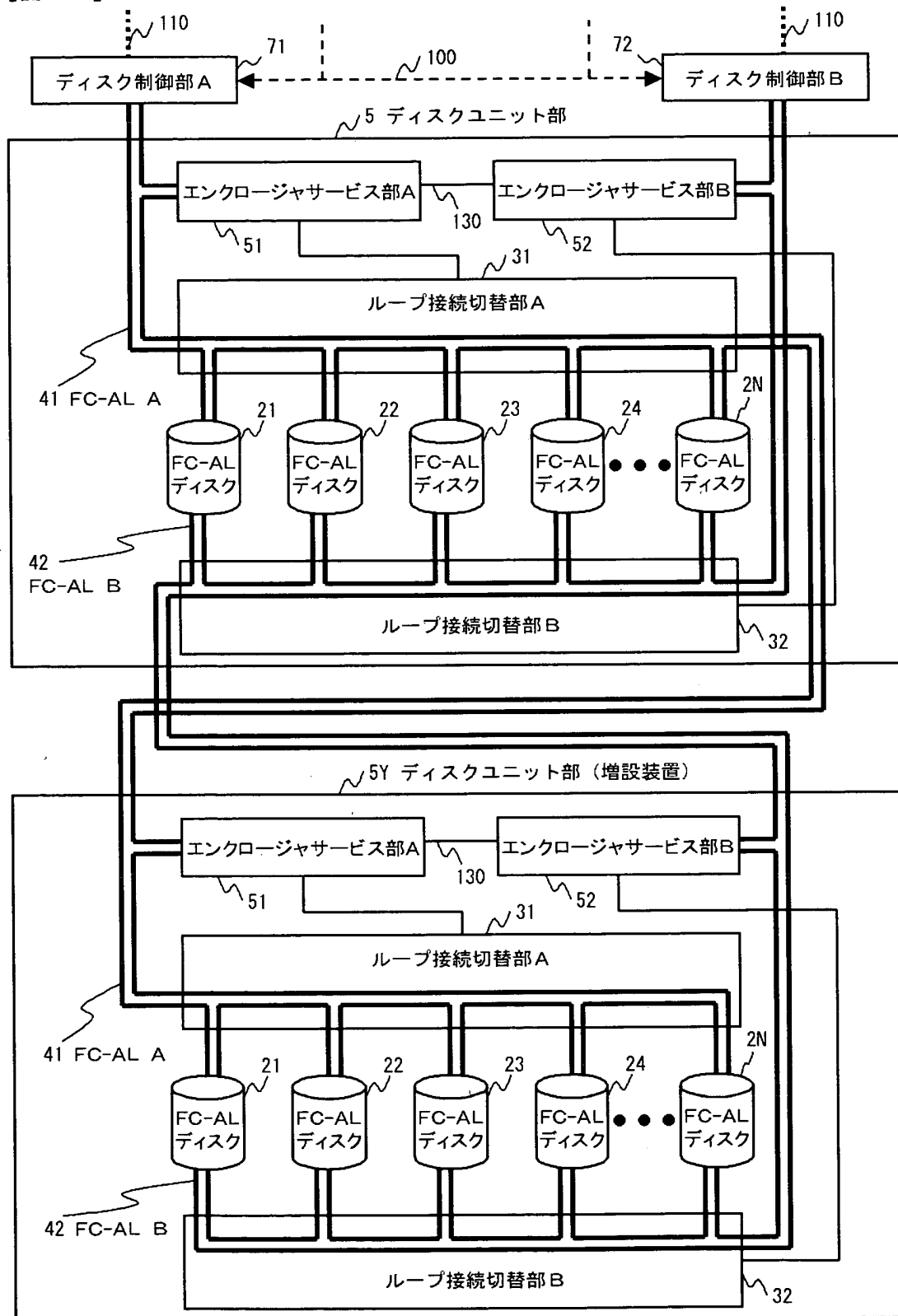
【図 9】

【図 9】



【図 10】

【図 10】



【書類名】 要約書**【要約】**

【課題】 2重化されたFC-A Lを有し且つその各々にエンクロージャサービス部が接続されたディスクアレイ装置において、全てのループで異常が発生しても、デバイスの切り離し、再接続を伴うループ診断が実施できるようにする。

【解決手段】 エンクロージャサービス部51、52はインタフェース130により相互に通信できる。ハートビート監視部511、521は、FC-A L41、42経由で送信されてくるハートビートコマンドが両ループで共に途絶えたと、双方のループで異常が発生したとして、FC-A L41、42に接続された全てのディスク21～2Nを切り離し、ループ異常を解消する。その後、ループ診断部711、721は連携して、FC-A L41、42へのディスクの再接続等を伴うループ診断を実施し、故障ディスクを特定し、故障デバイスを切り離すことによって、システムの運用を継続する。

【選択図】 図1

特願 2002-257545

出願人履歴情報

識別番号

[000004237]

1. 変更年月日

1990年 8月29日

[変更理由]

新規登録

住 所

東京都港区芝五丁目7番1号

氏 名

日本電気株式会社